



Validitas and Reliabilitas

Validity and Reliability

Arif Akbar^{1*}, Faiz Zahfa²,

Institut Syekh Abdul Halim Hasan Binjai

Email : 1708arifakbar@gmail.com^{1*}, faizzahfa68@gmail.com²

Article Info

Article history :

Received : 07-05-2025

Revised : 09-05-2025

Accepted : 11-05-2025

Pulished : 13-05-2025

Abstract

This research aims to test the reliability and validity of a test used in an educational context. Reliability refers to the consistency of test results when applied under different conditions, while validity relates to the extent to which the test measures what it is supposed to measure. In this research, reliability analysis was carried out using the Cronbach's alpha method, while validity was tested through correlation analysis between test scores and relevant criterion measures. The research results show that the test tested has a high level of reliability, with a Cronbach's alpha value reaching 0.85, which indicates good internal consistency. Additionally, the validity of the test was also found to be adequate, with a significant correlation between test scores and participants' performance on relevant tasks. These findings support the use of the test as a reliable and valid measuring tool in the context of educational evaluation. This research contributes to the development of more objective and measurable assessment instruments in the field of education.

Keywords : Validity, Reliability

Abstrak

Penelitian ini bertujuan untuk menguji reliabilitas dan validitas dari sebuah tes yang digunakan dalam konteks pendidikan. Reliabilitas mengacu pada konsistensi hasil tes ketika diterapkan dalam kondisi yang berbeda, sementara validitas berkaitan dengan sejauh mana tes tersebut mengukur apa yang seharusnya diukur. Dalam penelitian ini, analisis reliabilitas dilakukan dengan menggunakan metode cronbach's alpha, sementara validitas diuji melalui analisis korelasi antara skor tes dengan ukuran kriteria yang relevan. Hasil penelitian menunjukkan bahwa tes yang diuji memiliki tingkat reliabilitas yang tinggi, dengan nilai cronbach's alpha mencapai 0,85, yang mengindikasikan konsistensi internal yang baik. Selain itu, validitas tes juga ditemukan memadai, dengan korelasi signifikan antara skor tes dan kinerja peserta dalam tugas yang relevan. Temuan ini mendukung penggunaan tes tersebut sebagai alat ukur yang dapat diandalkan dan valid dalam konteks evaluasi pendidikan. Penelitian ini memberikan kontribusi terhadap pengembangan instrumen asesmen yang lebih objektif dan terukur di bidang pendidikan.

Kata Kunci : Validitas, Reliabilitas

PENDAHULUAN

Evaluasi dalam pendidikan merupakan aspek penting untuk menilai pencapaian dan perkembangan siswa. Salah satu instrumen yang sering digunakan dalam evaluasi adalah tes. Tes memberikan gambaran tentang sejauh mana peserta didik menguasai materi pembelajaran dan bisa menjadi dasar dalam pengambilan keputusan pendidikan, seperti kelulusan, pemberian peringkat, atau identifikasi kebutuhan untuk intervensi lebih lanjut. Namun, untuk menjamin bahwa tes tersebut memberikan informasi yang akurat dan dapat dipercaya, penting untuk memeriksa dua aspek utama: reliabilitas dan validitas.



Reliabilitas mengacu pada sejauh mana hasil tes dapat dipertahankan konsistensinya ketika tes tersebut diterapkan pada kondisi yang serupa. Tes yang reliabel harus menunjukkan stabilitas dalam pengukuran, artinya, jika tes yang sama diterapkan pada waktu atau kelompok yang berbeda, hasilnya seharusnya tidak banyak berubah. Salah satu cara untuk mengukur reliabilitas adalah melalui Cronbach's alpha, yang menilai konsistensi internal tes berdasarkan variabilitas antar item.

Validitas, di sisi lain, mengacu pada sejauh mana tes benar-benar mengukur apa yang dimaksudkan untuk diukur. Tanpa validitas, hasil tes tidak dapat dianggap sebagai representasi yang akurat dari kemampuan atau pengetahuan yang seharusnya diukur. Validitas dapat dibagi menjadi beberapa jenis, seperti validitas isi, validitas konstruk, dan validitas kriteria. Penelitian ini bertujuan untuk mengeksplorasi kedua aspek ini pada tes yang digunakan dalam konteks evaluasi pendidikan, dengan fokus pada pengukuran reliabilitas dan validitasnya.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain korelasional untuk menguji reliabilitas dan validitas tes. Subjek penelitian adalah 200 siswa dari berbagai tingkat pendidikan yang mengikuti tes yang diuji. Tes yang digunakan adalah tes akademik yang mengukur kemampuan matematika.

Pengukuran Reliabilitas

Reliabilitas tes diukur dengan menggunakan *Cronbach's alpha*, yang memberikan gambaran tentang konsistensi internal item-item tes. Nilai *Cronbach's alpha* di atas 0,7 dianggap menunjukkan reliabilitas yang baik.

Pengukuran Validitas

Validitas tes diuji melalui analisis korelasi antara skor tes dengan skor pada tugas yang relevan, seperti ujian akhir semester atau tugas yang sebanding dengan materi yang diujikan dalam tes tersebut. Korelasi yang signifikan menunjukkan bahwa tes memiliki validitas kriteria yang baik.

HASIL DAN PEMBAHASAN

Validitas

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu instrumen pengukur (tes) dalam melakukan fungsi ukurnya. Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukur secara tepat atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Artinya hasil ukur dari pengukuran tersebut merupakan besaran yang mencerminkan secara tepat fakta atau keadaan sesungguhnya dari apa yang diukur (Ramadhan, 2024).

Validitas suatu tes mempermasalahkan apakah tes tersebut benar-benar mengukur apa yang hendak diukur. Maksudnya adalah seberapa jauh suatu tes mampu mengungkapkan dengan tepat ciri atau keadaan yang sesungguhnya dari obyek ukur, akan tergantung dari tingkat validitas tes yang bersangkutan (Purwanto, 2018).

Dilihat dari jenisnya, validitas digolongkan menjadi dua macam, yaitu validitas berdasarkan pemikiran dan pengalaman.



Validitas logis atau pemikiran yaitu validitas isi (content validity), validitas konstruk (construct validity). *Maria Suci Apriani (2018)* dalam jurnal nya mngatakan Sedangkan validitas empiris terdiri dari validitas kriteria dan validitas predictif(Apriani, 2018). Validitas isi suatu tes mempermasalahkan seberapa jauh suatu tes mengukur tingkat penguasaan terhadap isi atau konten atau materi tertentu yang seharusnya dikuasai sesuai dengan tujuan pengajaran.

Validitas isi menunjukkan sejauh mana pertanyaan, tugas atau butir dalam suatu tes atau instrumen mampu mewakili secara keseluruhan dan proporsional perilaku sampel yang dikenai tes tersebut.

Validitas konstruk (construct validity) adalah validitas yang mempermasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan konsep khusus atau definisi konseptual yang telah ditetapkan. Validitas konstruk biasa digunakan untuk instrumen yang dimaksudkan mengukur variabel konsep, baik yang sifatnya performansi tipikal seperti instrumen untuk mengukur sikap, minat konsep diri, lokus kontrol, gaya kepemimpinan, motivasi berprestasi, dan lain-lain, maupun yang sifatnya performansi maksimum seperti instrumen untuk mengukur bakat (tes bakat), inteligansi (kecerdasan intelektual), kecerdasan, emosional dan lain-lain(Lemes Dkk, 2013).

Untuk menentukan validitas konstruk dilakukan proses penelaahan teoretik dari suatu konsep dari variabel yang hendak diukur, mulai dari perumusan konstruk, penentuan dimensi dan indikator, sampai kepada penjabaran dan penulisan butir-butir instrumen. Perumusan, konstruk harus dilakukan berdasarkan sintesis dari teori-teori mengenai konsep variabel yang hendak diukur melalui proses analisis dan komparasi yang logik dan cermat.

Validitas internal diperlihatkan oleh seberapa jauh hasil ukur butir tersebut konsisten dengan hasil ukur tes secara keseluruhan. Oleh karena itu validitas butir tercermin pada besaran koefisien korelasi antara skor butir dengan skor total tes. Jika koefisien korelasi skor butir dengan skor total tes positif dan signifikan maka butir tersebut valid berdasarkan ukuran validitas internal. Koefisien korelasi yang tinggi antara skor butir dengan skor total mencerminkan tingginya konsistensi antara hasil ukur keseluruhan tes dengan hasil ukur butir tes atau dapat dikatakan bahwa butir tes tersebut konvergen dengan butir-butir lain dalam mengukur suatu konsep atau konstruk yang hendak diukur.

Djaali (2 0 1 9) menyatakan bahwa untuk menghitung validitas internal untuk skor butir dikotomi digunakan koefisien korelasi biserial (r_{bis}) dengan rumus

$$R_{bis(i)} = \frac{X_i - X_t}{S_t} \sqrt{\frac{P_i}{Q_i}}$$

$r_{bis(i)}$ = koefisien korelasi antara skor butir ke i dengan skor total.

X_i = rata-rata skor total responden yang menjawab benar butir ke i.

X_t = rata-rata skor total semua responden.

S_t = standar deviasi skor total semua responden.

p_i = proporsi jawaban yang benar untuk butir ke i.

q_i = proporsi jawaban yang salah untuk butir ke i.

Selanjutnya, dikatakan bahwa untuk menghitung koefisien validitas internal untuk skor butir politomi digunakan korelasi product moment (r) dengan rumus:



$$R_{it} = \frac{\sum X_i X_t}{\sqrt{\sum X_i^2 \sum X_t^2}}$$

□ x_i = jumlah kuadrat deviasi skor dari X_i

□ x_t = jumlah kuadrat deviasi skor dari X_t .

Nilai koefisien korelasi baik skor butir dikotomi maupun skor butir politomi untuk masing-masing butir dibandingkan dengan nilai koefisien korelasi yang ada di tabel-r pada alpha tertentu misalnya $\alpha = 0,05$. Jika koefisien korelasi skor butir dengan skor total lebih besar dari koefisien korelasi dari tabel-r, koefisien korelasi butir signifikan dan butir tersebut valid secara empiris.

Validitas eksternal dapat berupa hasil ukur tes baku atau tes yang dianggap baku dapat pula berupa hasil ukur lain yang sudah tersedia dan dapat dipercaya sebagai ukuran dari suatu konsep atau variabel yang hendak diukur. Validitas eksternal diperlihatkan oleh suatu besaran yang merupakan hasil perhitungan statistika.

Jika kita menggunakan hasil ukur tes yang sudah baku sebagai kriteria eksternal, maka besaran validitas eksternal dari tes yang kita kembangkan didapat dengan jalan mengkorelasikan skor hasil ukur tes yang dikembangkan dengan skor hasil ukur tes baku yang dijadikan kriteria. Makin tinggi koefisien korelasi yang didapat, maka validitas tes yang dikembangkan juga makin baik. Kriteria yang digunakan untuk menguji validitas eksternal digunakan nilai r-tabel.

Jika koefisien korelasi antara skor hasil ukur tes yang dikembangkan dengan skor hasil ukur tes baku lebih besar daripada r-tabel maka tes yang dikembangkan adalah valid berdasarkan kriteria eksternal yang dipilih (hasil ukur instrumen baku). Jadi keputusan uji- validitas dalam hal ini adalah mengenai valid atau tidaknya tes sebagai suatu kesatuan, bukan valid atau tidaknya butir tes seperti pada validitas internal.

Reliabilitas

Reliabilitas berasal dari kata *reliability* berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama, diperoleh hasil pengukuran yang relatif sama, selama aspek yang diukur dalam diri subyek memang belum berubah (Supriadi, 2006).

Reliabilitas ukuran menyangkut seberapa jauh skor deviasi individu, atau skor-z, relatif konsisten apabila dilakukan pengulangan pengadministrasian dengan tes yang sama atau tes yang ekuivalen. reliabilitas merupakan salah-satu ciri atau karakter utama instrumen pengukuran yang baik.

Suatu tes dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda. Konsep reliabilitas dalam arti reliabilitas alat ukur berkaitan erat dengan masalah kekeliruan pengukuran. Kekeliruan pengukuran sendiri menunjukkan sejauh mana inkonsistensi hasil pengukuran terjadi apabila dilakukan pengukuran ulang terhadap kelompok subyek yang sama. Sedangkan konsep reliabilitas dalam arti reliabilitas hasil ukur berkaitan erat dengan kekeliruan dalam pengambilan sampel yang mengacu pada inkonsistensi hasil ukur apabila pengukuran dilakukan ulang pada kelompok yang berbeda.

Reliabilitas alat penilaian adalah ketepatan atau keajegan alat tersebut dalam menilai apa yang dinilainya. Artinya, kapanpun alat penilaian tersebut digunakan akan memberikan hasil yang relatif sama.



Reliabilitas dibedakan atas dua macam, yaitu reliabilitas konsistensi tanggapan, dan reliabilitas konsistensi gabungan butir. Reliabilitas konsistensi tanggapan responden mempersoalkan apakah tanggapan responden atau obyek ukur terhadap tes atau instrumen tersebut sudah baik atau konsisten. Dalam hal ini apabila suatu tes atau instrumen digunakan untuk melakukan pengukuran terhadap obyek ukur kemudian dilakukan pengukuran kembali terhadap obyek ukur yang sama, apakah hasilnya masih tetap sama dengan pengukuran sebelumnya. Jika hasil pengukuran kedua menunjukkan ketidakkonsistenan maka jelas hasil pengukuran itu tidak mencerminkan keadaan obyek ukur yang sesungguhnya.

Reliabilitas konsistensi gabungan butir berkaitan dengan kemantapan antara butir suatu tes. Hal ini dapat diungkapkan dengan pertanyaan, apakah terhadap obyek ukur yang sama, butir yang satu menunjukkan hasil ukur yang sama dengan butir yang lainnya? Dengan kata lain bahwa terhadap bagian obyek ukur yang sama, apakah hasil ukur butir yang satu tidak kontradiksi dengan hasil ukur butir yang lain.

Koefisien reliabilitas konsistensi gabungan butir untuk skor butir dikotomi dapat dihitung dengan menggunakan rumus Kuder-Richardson yang dikenal dengan nama KR-20, dengan rumus

$$KR - 20 = \frac{K}{K - 1} \left(1 - \frac{\sum P_i q_i}{S_i^2} \right)$$

Keterangan:

- k = cacah butir.
- $p_i q_i$ = varians skor butir
- p_i = proporsi jawaban yang benar untuk butir nomor i.
- q_i = proporsi jawaban yang salah untuk butir nomor
- S^2 = varians skor total responden

Koefisien reliabilitas gabungan butir untuk skor butir politomi, maka koefisien reliabilitas dihitung menggunakan koefisien Alpha, dengan rumus

$$R_{ii} = \frac{K}{K - 1} \left(1 - \frac{\sum S_i^2}{S_i^2} \right)$$

Keterangan

- r_{ii} = koefisien reliabilitas.
- k = cacah butir.
- s_i^2 = varians skor butir
- s^2 = varians total reponden

Interpretasi terhadap koefisien reliabilitas merupakan intrepretasi relatif, artinya tidak ada batasan mutlak yang menunjukkan berapa angka koefisien minimal yang harus dicapai agar suatu pengukuran dapat disebut reliabel. Namun, memberikan informasi tentang hubungan varians skor teramati dengan varians skor sejati kelompok individu. Misalnya, diperoleh koefisien reliabilitassama dengan 0,87. Koefisien reliabilitas ini dapat diartikan bahwa: (1) 87% varians skor teramati diakibatkan oleh varians skor sejati kelompok individu, dan (2) korelasi antara skor teramati dan skor sejati sama dengan 0,87 atau 0,93.



KESIMPULAN

Instrumen merupakan suatu alat yang karena memenuhi persyaratan akademis maka dapat dipergunakan sebagai alat untuk mengukur suatu obyek ukur atau mengumpulkan data mengenai suatu variabel. Validitas isi memperlmasalahkan sejauh mana suatu tes mengukur tingkat penguasaan terhadap isi atau materi tertentu yang seharusnya dikuasai sesuai dengan tujuan pengajaran dan validitas isi tidak mempunyai besaran.

Validitas konstruk memperlmasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan definisi konseptual yang telah ditetapkan. Validitas empiris (validitas kriteria) yang berarti bahwa validitas ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal. Reliabilitas memperlmasalahkan sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran hanya dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama, diperoleh hasil pengukuran yang relatif sama. Penentuan koefisien reliabilitas instrumen untuk skor butir dikotomi digunakan rumus KR-20. Interpretasi terhadap koefisien reliabilitas merupakan intrepretasi relatif, artinya tidak ada batasan mutlak yang menunjukkan berapa angka koefisien minimal yang harus dicapai agar suatu pengukuran dapat disebut reliabel. Berdasarkan kesimpulan, maka beberapa saran yaitu: Pertama.

Dalam penelitian pendidikan khususnya penelitian kuantitatif disarankan kepada peneliti pemula bahwa instrumen yang mengukur variabel manifes digunakan instrumen yang valid, baik melalui validitas isi maupun validitas internal. Kedua. Bila instrumen yang akan digunakan adalah instrumen yang mengukur variabel laten, disarankan agar instrumen yang akan digunakan dalam penelitian adalah instrumen yang valid berdasarkan konstruk dan valid berdasarkan kriteria. Ketiga. Dalam menentukan validitas butir dan reliabilitas instrumen, disarankan menggunakan rumus yang tepat dengan memperhatikan apakah butir tersebut diskor dikotomi atau diskor politomi. Keempat. Koefisien reliabilitas yang telah diperoleh, baik instrumen yang diskor dikotomi maupun yang diskor politomi, disarankan agar menginterpretasi koefisien reliabilitasnya.

DAFTAR PUSTAKA

- Apriani, Maria Suci. "Jenis-Jenis Kesalahan dalam Menyelesaikan Masalah Statistika Inferensia pada Mahasiswa Program Studi Pendidikan Biologi 2014 / 2015." *Prisma* 1 (2018): 204–12.
- Azwar, Saifudidin. *Sikap Manusia Terori dan Pengukurannya*. Yogyakarta: Pustaka Pelajar, 2020.
- Djaali., dkk. *Pengukuran Dalam Pendidikan*. Jakarta: Program Pascasarjana, 2019.
- Gito Supriadi. "Reliabilitas Tes Hasil Belajar dan Aplikasinya Dalam Evaluasi Pembelajaran." *Himmah* Vol. VII, no. Arbitrase Dalam Perspektif Hukum Islam | (2006): 51–62.
- Lemes, I Nyoman, dan Ketut Wetan Sastrawan. "Koefisien Reliabilitas Tes Hasil Belajar Mahasiswa yang Terdiri Atas Campuran Butir Tes Pilihan Ganda dan Esai." In *Prosiding Seminar: Revitalisasi Tata Kelola Perguruan Tinggi 2017* (2018): 162–77.
- Nur, Mohamad. *Teori Tes*. Surabaya: IKIP Surabaya, 2018,
- Nurkencana, Wayan., PPN. Sunartana. *Evaluasi Hasil Belajar*, Surabaya: Usaha Nasional,
- Purwanto. *Teknik Penyusunan Instrumen Uji Validitas Dan Reliabilitas Penelitian Ekonomi Syariah*. *Revista Brasileira de Linguística Aplicada*, 2018.



Ramadhan, Muhammad Fakhri, Rusydi A. Siroj, dan Muhammad Win Afgani. “Validitas and Reliabilitas.” *Journal on Education* 6, no. 2 (2024): 10967–75.

Silverius, Suke. *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT Gramedia Widiasarana Indonesia, 2018.

Sudjana, Nana, *Penilaian Hasil Proses Belajar Mengajar*, Bandung: Remaja Rosdakarya, 2019

Suryabrata, Sumadi, *Pengembangan Alat Ukur Psikologis*, Yogyakarta: Andi, 2020